

Learning from major accidents: graphical representation and analysis of multi-attribute events to enhance risk communication

R. Moura^{a,d}, M. Beer^{b,a,c}, E. Patelli^a & J. Lewis^a

^a*Institute for Risk and Uncertainty, University of Liverpool, United Kingdom*

^b*Institute for Computer Science in Civil Engineering, Leibniz University Hannover, Germany*

^c*Tongji University, Shanghai, China*

^d*National Agency for Petroleum, Natural Gas and Biofuels (ANP), Brazil*

ABSTRACT: Major accidents are complex, multi-attribute events, originated from the interactions between intricate systems, cutting-edge technologies and human factors. Usually, these interactions trigger very particular accident sequences, which are hard to predict but capable of producing exacerbated societal reactions and impair communication channels among stakeholders. Thus, the purpose of this work is to convert high-dimensional accident data into a convenient graphical alternative, in order to overcome barriers to communicate risk and enable stakeholders to fully understand and learn from major accidents. This paper first discusses contemporary views and biases related to human errors in major accidents. The second part applies an artificial neural network approach to a major accident dataset, to disclose common patterns and significant features. The complex data will be then translated into 2-D maps, generating graphical interfaces which will produce further insight into the conditions leading to accidents and support a novel and comprehensive “learning from accidents” experience.

1 Introduction

1.1 Perspectives on learning from accidents and understanding human errors

Major accidents have a multidimensional nature, arising from a wide range of contributing factors interacting in a seemingly random and sophisticated fashion to result in large-scale technological disasters. Many of these contributing factors are developed since the design conception, comprising of technical and non-technical issues and ultimately including the alluring influence of human errors.

The term “human error” has been coined in several different fields such as engineering, economics, psychology, design, management and sociology, with numerous interpretations and diverse objectives. Although most of the researchers and practitioners would (probably!) agree that human error can be generally understood as a failure to perform a certain task, the indiscriminate usage of the label “human error” to define some sort of human underperformance can be highly controversial. Hollnagel (1993, 1998), Woods et al. (2010) and Dekker (2014) claim that errors are best seen as a judgment in hindsight, or an attribution made about the behaviour of people after an event, being a quite misleading term, of limited practical use and nothing more than a tag. Conversely, Reason (1990, 2013) favoured the usage of the nomenclature, describing three necessary features to define human error: (i) plans; (ii) actions (or omissions); and (iii) consequences, surrounded by two situational factors: intention and absence of chance interference. Therefore, according to Reason, human errors can be acknowledged when an intention is reflected in a planned

sequence of actions which fails to accomplish its projected outcome, with observable consequences. The plan can be flawed or the action(s) can be imperfect, and some chance agency (e.g. an act of God or *force majeure*) is not recognisable.

The understanding of human error is typically encapsulated by a wider concept entitled “Human Factors”. This applied discipline has grown significantly after World War II, where the consideration of the human aspect was deemed necessary for achieving a realistic reliability assessment (Swain and Guttman, 1983; Dhillon, 1986). Since then, the multi-disciplinary nature of the human factors studies, which focus on the relationship between humans, tasks, technologies, organisations and the surrounding environment, allowed for some variances among the views and needs of engineers, psychologists, sociologists, managers and the general public. Hollnagel (1998) suggested that the original engineering and design approach aimed at analysing humans as components, in order to assign a human failure probability (or human error likelihood) to risk and safety assessments. Adopting a different perspective, psychologists attempted to understand mental processes and awareness mechanisms leading to erroneous actions, while sociologists were looking for flaws in the socio-technical system, usually attributing errors to management and organisational shortcomings.

But how do stakeholders see errors, especially after a major accident? Two common attitudes towards human errors were distinguished by Dekker (2014). The first one is what he called the “old view”, which considers errors as causes. On the other hand, the “new view” regards human errors as consequences of accidents, effects or symptoms of some sort of organisational shortcoming. Hollnagel (1998) highlighted that human error has been seen as the cause of events (when accidents are said to be due to the human intervention), the event itself (when an action, e.g. pressing the wrong button, is said to be a human error) or the consequences (the outcome of the action is said to be an error, e.g. the driver made the error of fuelling a petrol-fuelled car with diesel, inferring a car malfunction).

The general approach to human errors and the level of comprehension of human factors will deeply affect the process of learning from accidents. Major events are likely to trigger exacerbated societal reactions and impair communication channels, demanding an immediate and strong response from industry and authorities to ensure accountability. In those cases, acknowledging that humans are pivotal in any engineered system, the temptation to expose a scapegoat may limit the search to individuals who made errors throughout the lifecycle of the industrial process. If the concept of errors as causes of events prevails, the investigation will aim at the culprits’ exposure, the blame allocation and the imposition of penalties, and thus valuable lessons regarding organisational and technological aspects can be lost. Hopkins (2006) and Johnson (2008), discusses the tension between the requirement to learn as much as possible from events and the public pressure, highlighting that fear of litigation can act as an important barrier to learning the lessons from accidents.

That is why Reason’s characterisation of human errors turns out to be extremely beneficial, tying up two loose ends: firstly, the need for recognition and common understanding of human error, a deeply rooted concept in both technical and general public reality, serving as a useful bridge between the two worlds; and secondly, a clear and convenient definition, focused on the internal and external characteristics of the analysed subject and on the genesis of error. This approach allows for the search for profounder issues related to accidents and can help to reduce the knowledge gap

among authorities, the general public and wider stakeholder groups, in order to accomplish an improved learning environment.

1.2 Recent catastrophes in complex, multi-attribute accidents

The relevance of human factors and the impact of human errors in industrial accidents were extensively emphasised by contemporary studies. Human error was regarded as a major contributor to more than 70% of commercial airplane hull-loss accidents (Graeber, 1999). Correspondingly, according to Leveson (2004), operator errors can be considered the cause for 70–80% of accidents, given recurrent deviations between established practice and normative procedures. Considering the cost issue, a review by the United Kingdom Protection and Indemnity (P&I) Club indicated that US\$541 million per year is lost by the marine industry due to human errors (Dhillon, 2007).

Major accidents have the potential to capture the public's attention and demand strong responses from authorities. The Fukushima accident in March 2011 has triggered a considerable shift in the way the nuclear industry was seen by governments and the general public.

Mr. Kiyoshi Kurokawa, the chairman of the independent investigation commission for the Fukushima nuclear accident, stated in the official report that the tsunami and technical issues were not the sole reasons for the tragedy, declaring that human factors as well as deeper Japanese cultural issues were vital contributors (Kurokawa et al, 2012): *"What must be admitted – very painfully – is that this was a disaster 'Made in Japan'. Its fundamental causes are to be found in the ingrained conventions of Japanese culture (...). Therefore, we conclude that the accident was clearly 'man-made'"*.

This wide-reaching accident was drawn into the international media's spotlight. In a global perspective, it has driven the German government to immediately shut-down eight reactors (their restart is highly unlikely) and reject the construction of new units (Schneider et al., 2012). Thus, the phasing-out of nuclear plants and the replacement with renewables or other energy sources in Germany seems to be irreversible. Also, the former Republic of China's president Ma Ying-Jeou decided to limit the operating lifespan of nuclear power plants to 40 years, and declared that the continuation of the ongoing construction of a new nuclear power unit in Taiwan would be decided by a public referendum (Ishikawa, 2015). It appears that the political climate due to the public's perception of risk is highly unfavourable to nuclear power, especially after Fukushima.

Recently, prosecutors called for the death penalty for the captain of the MV Sewol, a South Korean ferry that sank in April 2014 and left 304 fatalities, most of them pupils on a school trip. He was found guilty and sentenced to 36 years in prison, and 14 crewmembers were jailed from 9 to 25 years. Families and protesters affirmed that lessons were not learnt from a series of previous accidents (BBC, 2014).

These are examples, to name but a few, of how the process of learning from accidents can be compromised if risks are not adequately communicated, due to the natural distance between experts' views and the public's perception.

From an engineering perspective, the highly complex interaction between operators, technology and organisations is a recurring subject arising from investigations involving major events. The *Bureau d'Enquêtes et d'Analyses pour la sécurité de l'aviation civile* (2011) official report on the AF-447 Rio-Paris Airbus A-330 accident on 1st June 2009 acknowledged an apparently simple equipment defect (icing of the Pitot probes) resulting from a design failure, which led to some inconsistencies of the flight speed indicators. This deficiency triggered several human-related events (wrong system diagnosis and inappropriate control inputs, among others), ultimately resulting in the airplane hull-loss in the Atlantic Ocean, with 228 victims. The investigation report also highlighted some intricate factors such as the de-structuring of the task-sharing in the cockpit during the response to the anomalous event, training shortcomings in a predictable flight mode (manual handling of the airplane in high altitudes), and the lack of indication of the airspeed inconsistencies in the flight console, exposing an complex combination of several factors leading to the catastrophe.

These major accident examples, mostly involving up-to-date technologies with numerous systems under normal operation (e.g. Airbus (2016) states that around 1,200 A330 airplanes are operated by over 100 companies, meaning that an aircraft takes off or lands somewhere every 20 seconds!), illustrates the complexity behind erroneous actions, mental models, technology, organisational issues, culture and the environment in high-technology industries. This highly interdisciplinary and intricate setting, including the influence of the utmost public and media attention and the fact that human errors are palpable and a compelling argumentation to explain undesirable events, brings a substantial challenge to stakeholders. How to develop means to learn from multi-attribute events and translate these lessons into an approachable scheme for researchers, practitioners, policymakers and society, in order to communicate and tackle risks appropriately?

Consequently, this work aims to overcome barriers to dealing with complex datasets containing incident/accident information, by means of applying an unsupervised learning neural network approach to a proprietary accident dataset. The Multi-Attribute Technological Accidents Dataset (MATA-D) presented by Moura et al. (2016) will be converted into two-dimensional graphical representations of accidents and their corresponding surrounding factors. The 2-D interfaces will provide innovative means to communicate high-technology risks and to disclose tendencies that could lead to the genesis of errors, facilitating and enhancing interactions among internal stakeholders and the general public.

2 Analysis Method

2.1 Capturing the complexity underlying major accidents: fit for the past, adaptable to the future

The underlying dynamics observed during critical events is so great that some renowned accident causation theorists consider the failures in complex, tightly coupled systems as inevitable (Perrow, 1999), or not prospectively foreseeable (Taleb, 2007). This is due to the acknowledged difficulties to capture and understand all facets of socio-technical systems and all circumstances leading to catastrophes, which pose a challenge to researchers and practitioners. As a result, any method to capture lessons from accidents will have inherent limitations. In this work, a data-based approach which starts from the available information (detailed accounts from major accidents) and uses an artificial neural network process to generate useful knowledge is proposed. The dataset structure is

flexible and expandable: new accidents can be added, or prospective analyses can be conducted in order to increase the database.

Other dataset frameworks inspired by novel accidents causality models based on systems theory and system thinking, such as the System-Theoretic Accident Model and Processes (Leveson, 2011), could be used for accident analyses, and the artificial neural network approach presented in this work would be equally applicable. However, the data required to construct a reasonable model based on some new approaches would require accident investigators to have previous knowledge of those approaches and carry out a targeted data collected process to suit the framework. Consequently, it is unlikely that the massive amount of data from early major accident accounts currently available would be immediately adaptable. Even considering a broader, virtually ideal holistic approach, potential influencing factors such as the percentage of profit allocated to safety, the influence of political decisions on industrial segments or the impact of the change of controllers and investors on companies, would not be fully available for earlier accidents, implying a new start point for data collection and thus a drawback to embrace past events.

Therefore, the choice of the dataset framework took into account the possibility of starting from the available information reflected on past accounts to generate immediate knowledge. On the contrary, to construct a dataset based on a fully new framework, many years of catastrophic events (fortunately major accidents are rare) would be necessary, and even the most comprehensive model would still hold intrinsic limitations.

Additionally, this study presents an expansion of conventional reductionist models. The focus of the analysis method lies on the examination of the interfaces among contributing factors, instead of specifying root-causes in a classical chain of directly related events. Although accident reports are intended to present a logical explanation to accidents, usually comprising a sequence of events through time, non-linear interactions among contributing factors can be also identified during in-depth investigations, allowing a systemic learning process. Irrespective from the accident causation method used to understand events, it was possible to classify all of them under the common framework in which the dataset was based, as seen on Moura *et al.* (2016).

For example, The Piper Alpha Accident Report (Cullen, 1990) revealed design, construction, management, operational and human factors interfacing in an undesirable fashion to result in 167 fatalities and billions of pounds in property losses in July 1988. The recommendations (106 in total) arising from the report addressed changes to oil & gas offshore facilities, industry, the UK government and trade unions. New legislation (The Safety Case Regulations) was developed after a full review of existing legislative arrangements, progressing from the former prescriptive regime to the current performance-based safety management model. Even the responsibility for safety oversight was transferred from the Department of Energy which used to regulate both revenue collection and safety, to the Health & Safety Executive (Pate-Cornell, 1988).

Therefore, accident narratives arising from detailed investigation processes can give impetus to broader safety improvement measures, and the collective understanding of previous occurrences is able to reveal new interfaces, contribute to a holistic safety understanding and improve stakeholders' risk awareness level.

2.2 The Multi-attribute Technological Accidents Dataset (MATA-D)

Many researchers (e.g. Swain, 1990; International Atomic Energy Agency, 1990; Grabowski, 2009) referred to the lack of reliable data on human performance in high-technology systems and the complications associated with the collection, consistency and interpretation of data.

In addition, most of the near-misses datasets contains condensed descriptions of events, generally limiting the information to direct or immediate causes (e.g. operator failure, equipment defect), due to usual constraints (i.e. time, budget, human resources) to conduct in-depth analysis of inconsequential events. The main shortcoming with non-detailed data is that the context in which workers are placed is usually overlooked. However, it is widely accepted (Reason, 1990; Hollnagel, 1998; Cooper, 1996; Strater, 2000; Deker, 2014) that the context is actually the central element to be studied, in order to provide a full picture and a better understanding of undesirable events.

The European Safety, Reliability and Data Association (2015) stated that major accident investigations allow for a detailed analysis of preventive and protective systems, as well as the exploration of events and surroundings conditions leading to accidents. In addition, high-impact events usually provide impetus for the application of lessons learned to minimise reoccurrence, as observed in the wake of disasters such as the Texas City Refinery, the offshore production platform Piper Alpha and the Nuclear Power Plant in Fukushima (Fukasawa, 2012; Dahle, 2012). Society's risk perception, the regulatory approach and industry's behaviour towards safety were affected in a global scale by these events.

Although major accidents seem to be one of the finest sources of information available, they are considered to be rare events (Reason, 1997; Taleb, 2007), and the currently available data might not be enough for the application of traditional statistical approaches. Thus, a method to allow the seamless learning process between different industrial sectors is necessary, in order to generate sufficient data for a suitable analysis. Nevertheless, transversal learning is not trivial, due to differences among technologies, industrial jargon and contexts in complex systems. To overcome these issues, the authors developed a major-accident dataset, using a common framework, the Contextual Control Model used as basis for the Cognitive Reliability and Error Analysis Method (Hollnagel, 1998), to classify data captured from investigation reports prepared by regulators, investigation commissions, government bodies, insurance companies and industry experts to explain the contributing factors and causes behind major accidents. The framework is comprehensive, containing 53 contributing factors distributed in groups (man, technology and organisation) and subgroups (erroneous actions, specific cognitive functions, temporary and permanent person-related functions; equipment, procedures, temporary and permanent interface; communication, organisation, training, ambient and working conditions). The data structure is represented in Figures 1, 2 and 3. Further information about the decision to develop a new dataset and all the details regarding the creation and content of the Multi-attribute Technological Accidents Dataset (MATA-D) can be found in Moura et al. (2016).

The use of high-consequence accident reports to feed the MATA-D proved to have many benefits. Deep investigations involve internal and external experts in the search for evidence and to disclose contributing factors and relevant interactions between humans, technology and organisations.

Therefore, uncertainties associated with the consistency of the input data were reduced, and the selected framework permitted the classification of events from diverse industrial backgrounds under a common taxonomy, making them comparable. Consequently, the MATA-D structure allows for the application of mathematical methods, aiming at the disclosure of common patterns and at the recognition of significant features. This way, the genesis of multi-attribute events can be better understood and communicated.

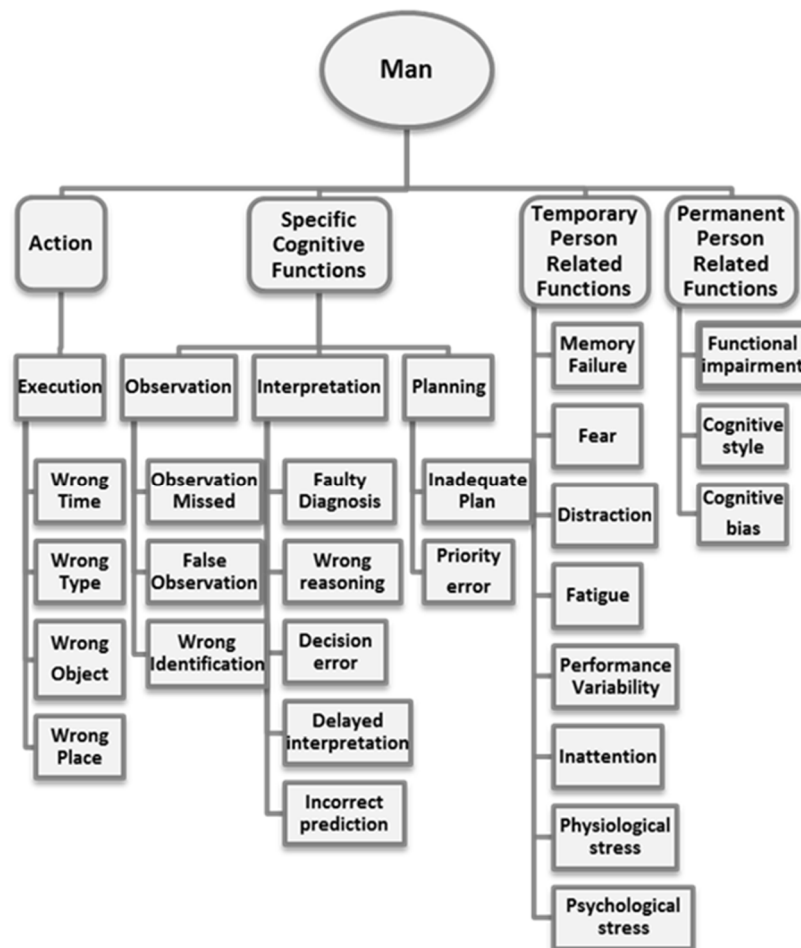


Figure 1. "Man" categorisation, adapted from Hollnagel (1998)

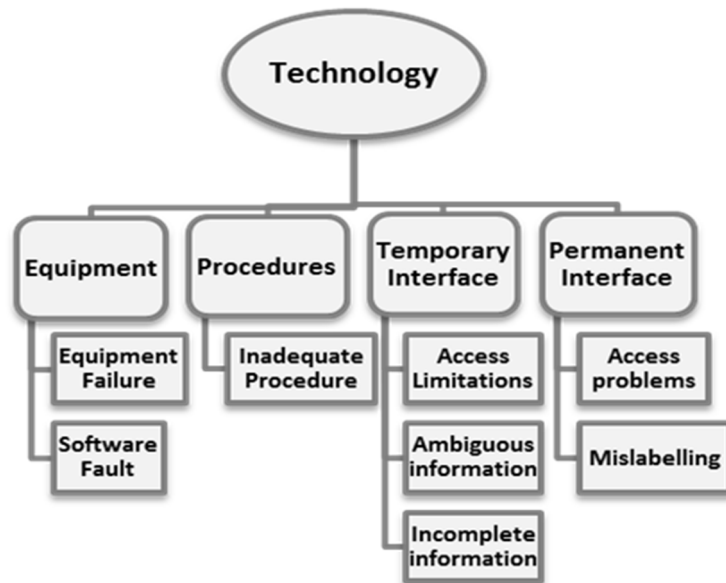


Figure 2. “Technology” categorisation, adapted from Hollnagel (1998)



Figure 3. “Organisation” categorisation, adapted from Hollnagel (1998)

The dataset includes 238 major accidents from several industrial segments, as detailed in Table 1.

Table 1. MATA-D Accidents distribution by industrial sector, after Moura et al. (2016)

Industrial Sector	Accidents	
	#	%
Refinery	39	16.39
Upstream (oil & gas)	37	15.55
Chemicals	29	12.18
Petrochemicals	25	10.50
Nuclear Power Plants	23	09.66
Civil Construction	16	06.72
Terminals & Distribution	15	06.30
Aviation	13	05.46
Gas Processing	09	03.78
Metallurgical industry	07	02.94
Waste Treatment Plant	05	02.10
Food Industry	04	01.68
Others	16	06.72

2.3 The data mining process

The aim of the data mining process is to disclose common structures among accidents and significant features within the major-accident dataset. In this work, making an attempt to go beyond the general statistical analysis presented by Moura et al. (2016), a well-known clustering approach named Self-Organising (or Kohonen) Maps (SOM), developed by Kohonen (1998), is applied to the MATA-D. The objective is to convert the 53-dimensional dataset (a matrix of 238 accidents holding 53 possible contributing factors each) into a low-dimensional (i.e. 2-D) array, enabling the data visualisation and interpretation.

Essentially, the SOM algorithm consists of an initialisation followed by three processes: (i) competition; (ii) cooperation; and (iii) adaptation. The network learning begins with the attribution of arbitrary values for the initial weight vectors. Then, the training starts with the competition process, where the winning output neuron (best matching node) is the one which minimises the Euclidean Distance $\|\vec{x} - \vec{m}_i\|$ for each input pattern (Eq. (1)). After defining the output winner, the cooperation process consists of the application of a neighbourhood function (usually the Gaussian function Eq. (2)) to define the spatial influence of the best matching unit upon the neighbour neurons. The last process is the adaptation one, where the weights of all neighbour neurons are sequentially updated while both the learning rate and the neighbourhood function decrease with time, following the Eq. (3). This sequence is repeated (through iterations) until the map converges (Kohonen, 2001; Andreev and Argyrou, 2011).

$$v(t) = \arg \min_{i \in \Omega} \|x(t) - m_i(t)\| \quad (1)$$

$$h_{ci}(t) = \alpha(t) \cdot \exp\left(\frac{\|r_c - r_i\|^2}{2\sigma^2(t)}\right) \quad (2)$$

$$m(t + 1) = m_i(t) + \alpha(t) h_{ci}(t) [x(t) - m_i(t)] \quad (3)$$

Kohonen (2013) has shown that a variation of the updating rule (Eq. 3) would be useful to eliminate convergence complications and generate steadier asymptotic m_i values. Therefore, a batch-learning version of SOM (Batch-SOM) was revealed for practical applications, in order to generate more consistent outcomes. Eq. (4) shows the Batch-SOM updating rule developed by Kohonen.

$$m_i^* = \frac{\sum_j n_j h_{ji} \bar{x}_j}{\sum_j n_j h_{ji}} \quad (4)$$

With the new update rule (Eq. 4), the definition of a learning parameter α is no longer necessary, as the batch-learning implies that the codevectors are being updated once, rather than in a recursive fashion. Each best matching node m_i^* represents the centroid of an influence region defined by \bar{x}_j (the mean value of a group of input vectors $x(t)$), the neighbourhood function h_{ji} and the number n_j of samples.

The representation of the dataset (a 53-dimension input space, Figure 4) in a 2-D topographic map shows the MATA-D events organised by similarity, i.e. accidents with analogous contributing factors will be close to each other (e.g. Figure 5). This enables the generation of clusters which can be analysed in an integrated way, revealing tendencies within a group of major accidents.

One of the most important features of the SOM's learning process is that it comprises an unsupervised learning process, dismissing the need for any pre-classification, pre-selection of the number of clusters or the definition of main/leading factors (Kohonen et al, 1996). Consequently, the data mining process is not affected by external parameters, avoiding potentially biased concepts regarding the main factors influencing accidents and potentially leading to human errors.

3 Results

3.1. Clustering Results

The application of the SOM algorithm brought together accidents by resemblance – the more similar the accidents are, the closer they are positioned in the output space. Figures 4 and 5 shows the input space (a matrix 238 x 53) and the output space (a 2-D representation of the events).

$$\begin{matrix} A_{1.1} & \dots & A_{1.53} \\ \vdots & \ddots & \vdots \\ A_{238.1} & \dots & A_{238.53} \end{matrix}$$

Figure 4. Input Space: A dataset containing 238 samples x 53 features

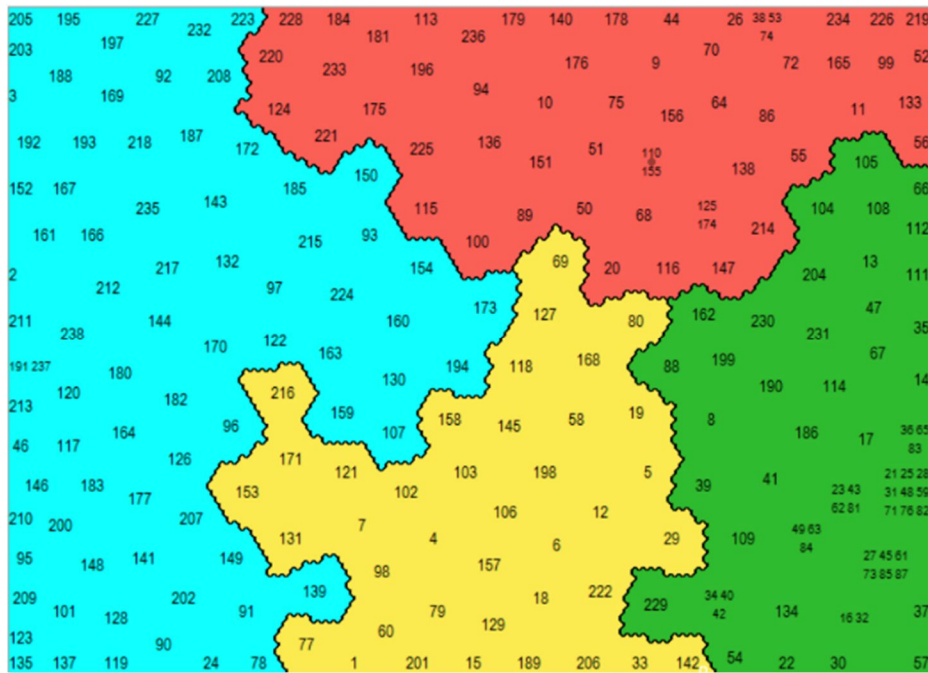


Figure 5. Output Space: A reorganised 2-D grid

Maps were generated from the 238 records with 53 possible attributes each. The output space was trained with 111 batches, and the width of the kernel (the radius of the Gaussian neighbourhood - Eq. (2)) was 0.41. This value is set to be the smallest to form meaningful clusters while maintaining the best possible representation of data differences, with attribute values being averaged less. Maps were produced by the expert version of Viscovery SOMine® software, in order to enhance graphical visualisation.

The Viscovery® SOMine software has a clustering quality indicator, a histogram which classifies conceivable groupings by attributing an index for each possible clustering arrangement. The 4-cluster final map attained the highest quality measure (Figure 6) and thus was adopted as a useful arrangement for further interpretation.

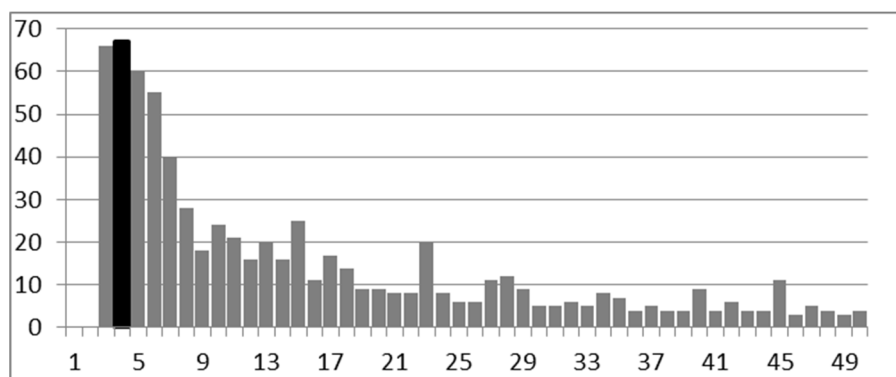


Figure 6. Cluster Quality Indicator (Viscovery SOMine®)

Other analysis methods can be used to evaluate the quality of the map, such as the error quantisation and frequency computation. For the current application, quantisation errors were very small, and both the data frequency and the error quantisation were evenly distributed throughout the map, indicating that the network was well-trained and the mapping quality is good.

Tables 2 and 3 present the clusters' statistical results.

Table 2. Clusters' Features

Cluster	Events #	Contributing Factors #				
		Min	Max	Mean	Median	Mode
C1	80	04	24	9.62	09	09
C2	57	01	10	4.56	04	02
C3	39	05	22	8.92	08	08
C4	62	01	06	3.10	03	02

Table 3. Overall dataset statistics (After Moura et al., 2016) and new clustering statistics

	Accidents #	Overall 238	C 1 80	C 2 57	C 3 39	C 4 62
Erroneous Actions	Wrong Time	14.7%	13.8%	10.5%	41.0%	3.2%
	Wrong Type	11.8%	11.3%	7.0%	30.8%	4.8%
	Wrong Object	2.5%	3.7%	3.5%	2.6%	0.0%
	Wrong Place	31.5%	52.5%	36.8%	12.8%	11.3%
Observation	Observation Missed	15.5%	20.0%	12.3%	23.1%	8.1%
	False Observation	3.4%	6.3%	3.5%	0.0%	1.6%
	Wrong Identification	2.5%	5.0%	0.0%	5.1%	0.0%
Interpretation	Faulty diagnosis	13.0%	26.3%	8.8%	12.8%	0.0%
	Wrong reasoning	11.3%	20.0%	1.8%	25.6%	0.0%
	Decision error	9.2%	5.0%	17.5%	17.9%	1.6%
	Delayed interpretation	4.6%	8.7%	1.8%	7.7%	0.0%
	Incorrect prediction	3.8%	7.5%	1.8%	2.6%	1.6%
Planning	Inadequate plan	23.0%	10.0%	7.0%	25.6%	1.6%
	Priority error	7.1%	6.3%	8.8%	15.4%	1.6%
Temporary Person-related Functions	Memory failure	0.9%	1.3%	1.8%	0.0%	0.0%
	Fear	2.1%	1.3%	0.0%	5.1%	3.2%
	Distraction	5.9%	11.3%	3.5%	7.7%	0.0%
	Fatigue	2.9%	7.5%	0.0%	2.6%	0.0%
	Performance Variability	1.4%	5.0%	1.8%	0.0%	0.0%
	Inattention	2.1%	2.5%	0.0%	5.1%	1.6%
	Physiological stress	0.8%	1.3%	1.8%	0.0%	0.0%
	Psychological stress	2.9%	5.0%	1.8%	2.6%	1.6%
Permanent Person-related Functions	Functional impairment	0.4%	0.0%	0.0%	2.6%	0.0%
	Cognitive Style	0.0%	0.0%	0.0%	0.0%	0.0%
	Cognitive bias	7.1%	15.0%	1.8%	10.3%	0.0%
Equipment Failure	Equipment failure	55.0%	33.8%	22.8%	94.9%	87.1%
	Software fault	2.5%	6.3%	0.0%	2.6%	0.0%
Procedures	Inadequate procedure	44.1%	78.7%	42.1%	38.5%	4.8%
Temporary Interface Problems	Access limitations	1.3%	3.7%	0.0%	0.0%	0.0%
	Ambiguous information	2.5%	5.0%	0.0%	5.1%	0.0%
	Incomplete information	17.6%	36.2%	7.0%	20.5%	1.6%
Permanent Interface Problems	Access problems	1.7%	3.7%	0.0%	2.6%	0.0%
	Mislabelling	1.7%	2.5%	1.8%	0.0%	1.6%
Communication	Communication failure	10.5%	16.3%	5.3%	20.5%	1.6%
	Missing information	20.6%	37.5%	14.0%	15.4%	8.1%
Organisation	Maintenance failure	34.9%	56.3%	14.0%	33.3%	27.4%
	Inadequate quality control	60.5%	81.3%	24.6%	79.5%	56.5%
	Management problem	9.2%	12.5%	5.3%	23.1%	0.0%
	Design failure	66.0%	85.0%	50.9%	87.2%	41.9%
	Inadequate task allocation	60.1%	95.0%	68.4%	48.7%	14.5%
	Social pressure	7.1%	17.5%	3.5%	0.0%	1.6%
Training	Insufficient skills	36.1%	56.3%	12.3%	76.9%	6.5%
	Insufficient knowledge	35.3%	60.0%	17.5%	56.4%	6.5%

Ambient Conditions	Temperature	1.3%	1.3%	0.0%	2.6%	1.6%
	Sound	0.0%	0.0%	0.0%	0.0%	0.0%
	Humidity	0.0%	0.0%	0.0%	0.0%	0.0%
	Illumination	0.8%	1.3%	1.8%	0.0%	0.0%
	Other	0.0%	0.0%	0.0%	0.0%	0.0%
	Adverse ambient conditions	7.1%	2.5%	14.0%	10.3%	4.8%
Working Conditions	Excessive demand	5.5%	6.3%	8.8%	5.1%	1.6%
	Inadequate work place layout	2.5%	1.3%	7.0%	2.6%	0.0%
	Inadequate team support	3.4%	6.3%	0.0%	7.7%	0.0%
	Irregular working hours	3.8%	10.0%	1.8%	0.0%	0.0%

3.2 Clusters Description

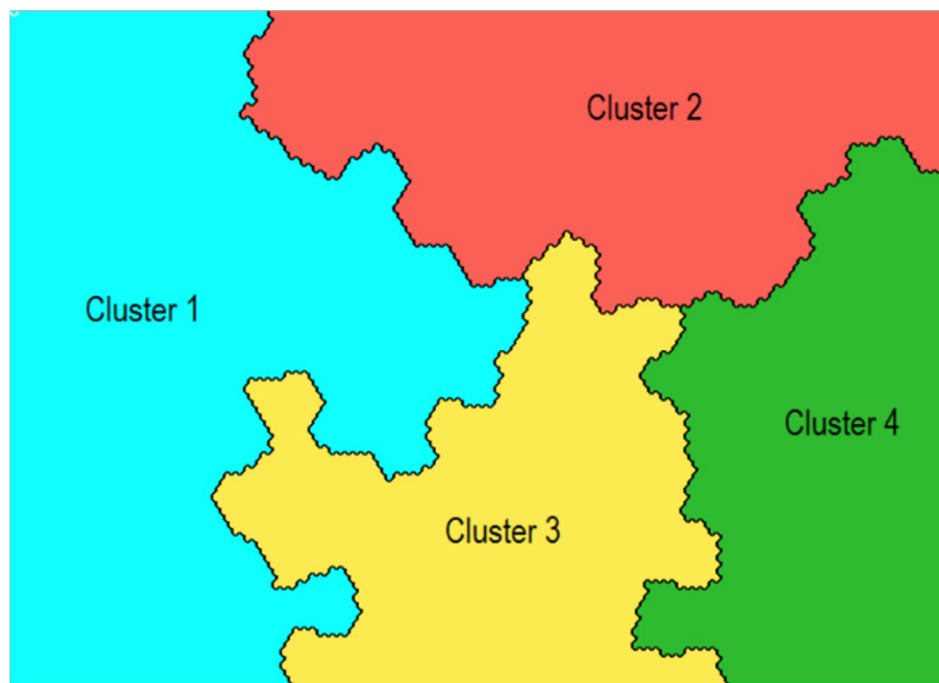


Figure 7. Clusters Identification

Figure 7 presents the accidents positioned by similarity in the grid, separated in 4 (four) clusters with different dominating characteristics.

Cluster 1 is the largest single group, covering 80 accidents which encloses between 4 and 24 contributing factors. It is largely dominated by Inadequate Task Allocation (95.0%), when the organisation of work is lacking due to deficient scheduling, task planning or poor rules and principles. Additionally, accidents within this cluster were deeply influenced by Design Failure (85.0%) and Inadequate Quality Control (81.3%). From an organisational perspective, factors such as Insufficient Knowledge (60.0%), Maintenance Failure (56.6%) and Missing Information (37.5%) were also significant. The most important technological contributor was the Inadequate Procedure factor with 78.7% of incidence, followed by Incomplete Information (36.2%). From a human factors perspective, erroneous actions labelled as Wrong Place (when actions in a planned sequence are

omitted/skipped, repeated, reversed or when an unnecessary action is taken) were in 52.5% of the cluster, mainly accompanied by interpretation issues, specially Faulty Diagnosis, with 26.3% of incidence. Observation Missed and Wrong Reasoning were also noteworthy human-related factors, both appearing in 20% of the cluster.

Cluster 2 grouped 57 accidents, all varying from 1 to 10 contributing factors and with low mean, median and mode figures (Table 2). Organisational factors such as Inadequate Task Allocation (68.4%) and Design Failure (50.9%), as well as the technological factor Inadequate Procedure (42.1%) were the most frequent in this grouping. Some hostile ambient and working conditions were highlighted, with Adverse Ambient Conditions (14.0%) and Inadequate Workplace Layout (7.0%) standing above the overall data distribution. Decision Error (17.5%) was a noticeable human contributor, highlighting cases where workers were unable to make a decision or have made the wrong choice among possible alternatives.

The leading factor for Cluster 3, which contains 39 major accidents, is a technological aspect labelled Equipment Failure, populating 94.9% of the grouping area. This cluster also presented very strong organisational factors, as Design Failure (87.2%), Insufficient Skills (76.9%), Management Problem (23.10%) and Communication Failure (20.5%) attained their maximum values in this cluster, being also the Inadequate Quality Control factor very relevant, with 79.5% of incidence. The human factors incidence is quite substantial, with actions occurring at the wrong time (41.0%) or being of the wrong type (30.8%). These cases include omitted, premature or delayed actions, as well as using disproportionate force, magnitude, speed or moving in the wrong direction. These human erroneous actions were accompanied by all three levels of cognitive functions, i.e. observation, represented by Observation Missed (23.1%), interpretation, with Wrong Reasoning (25.6%) and Decision Error (17.9%); and planning, with both Inadequate Plan (25.6%) and Priority Error (15.4%) attaining their maximum incidence. It is worth to notice that the number of contributing factors for each event in this cluster fluctuated from 5 to 22, with a mean of approximately 9 and median and mode of 8.

Cluster 4 contains 62 events, each one encompassing 1 to 6 contributing factors, recording a mean of approximately 3, median of 3 and a mode of 2 features. For most of the accidents in this cluster (i.e. 87.1%), an Equipment Failure was the most frequent contributor to accidents, followed by Inadequate Quality Control (56.5%), Design Failure (41.9%) and Maintenance Failure (41.9%).

Figures 8 to 21 show the self-organising maps for individual features. These figures detail how individual characteristics were distributed in the map after the application of the SOM algorithm. Colder colours (tending to blue) mean the absence of a feature, while warmer colours (tending to red) mean the presence of a contributing factor. Multiple intersections of warm colours in different individual SOM maps can be interpreted as an interface/relationship.

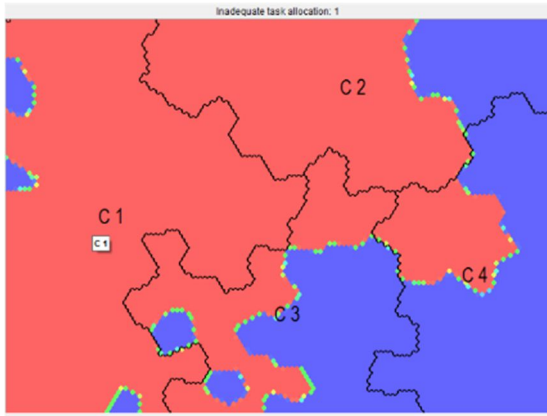


Figure 8. Inadequate Task Allocation SOM

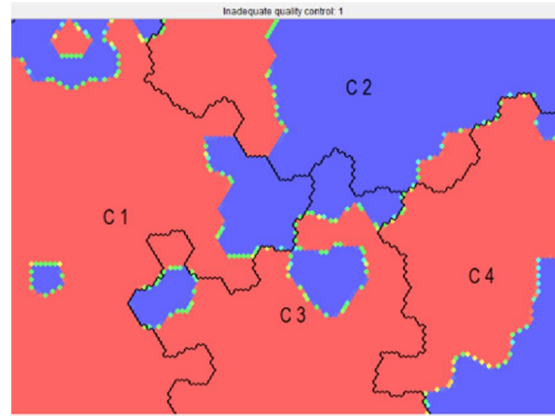


Figure 9. Inadequate Quality Control SOM

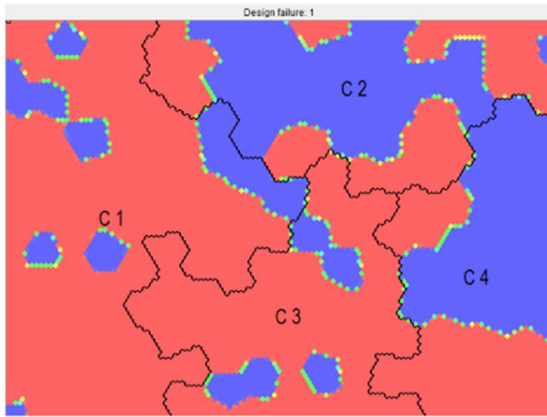


Figure 10. Design Failure SOM

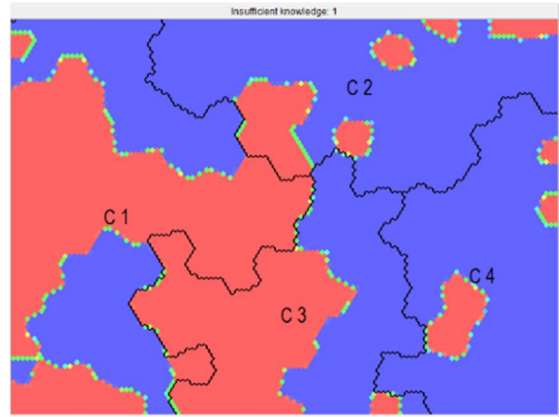


Figure 11. Insufficient Knowledge SOM

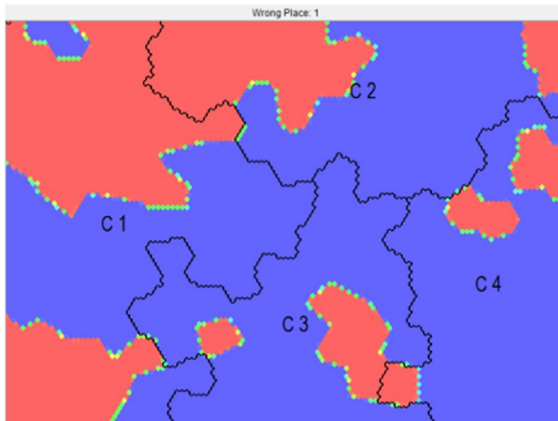


Figure 12. Wrong Place SOM

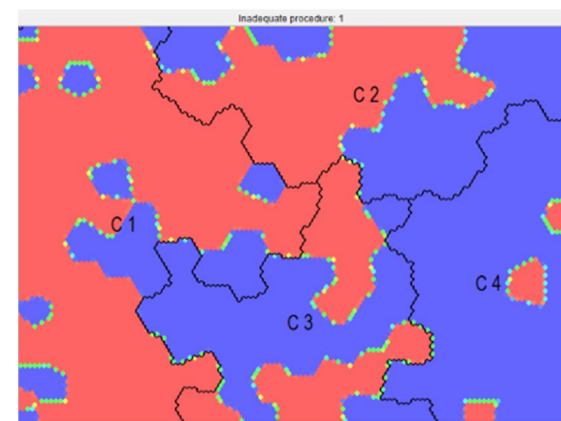


Figure 13. Inadequate Procedure SOM

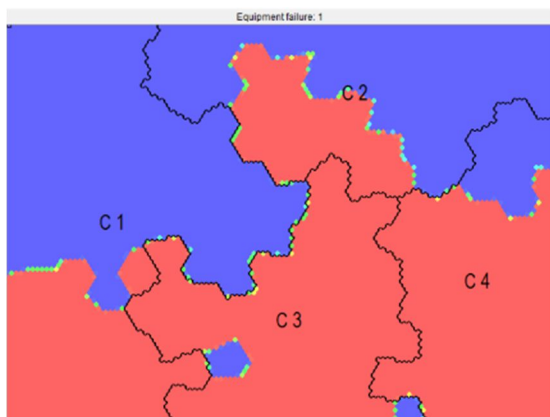


Figure 14. Equipment Failure SOM

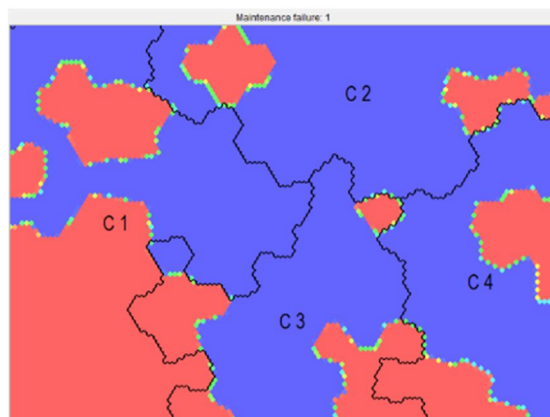


Figure 15. Maintenance Failure SOM

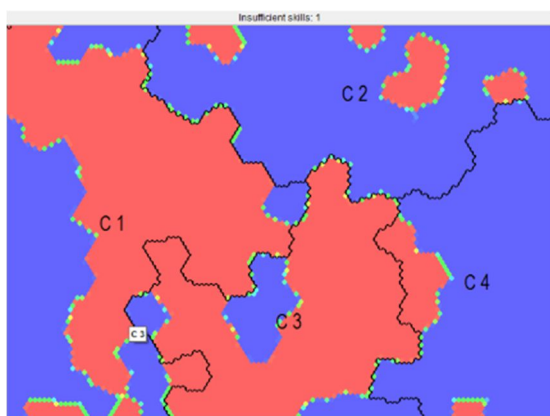


Figure 16. Insufficient Skills SOM

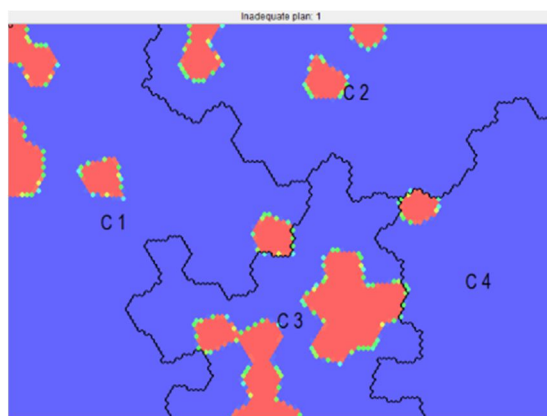


Figure 17. Inadequate Plan SOM

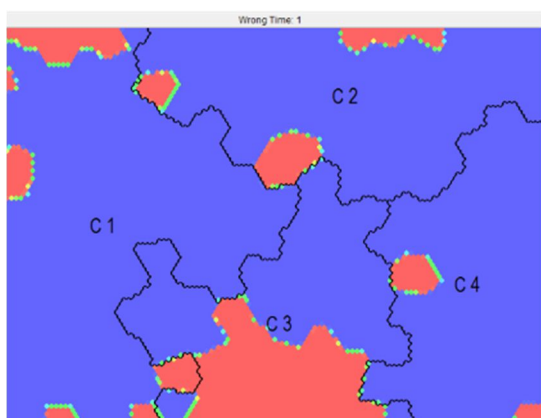


Figure 18. Wrong Time SOM

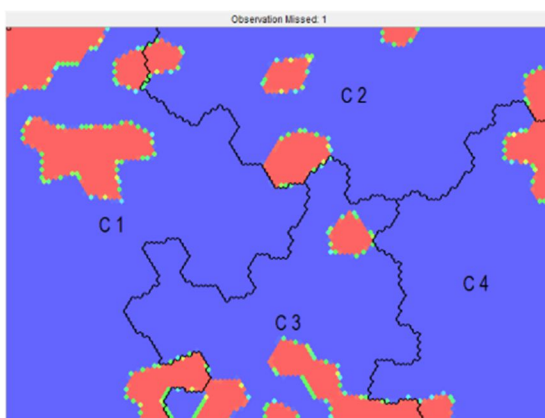


Figure 19. Observation Missed SOM

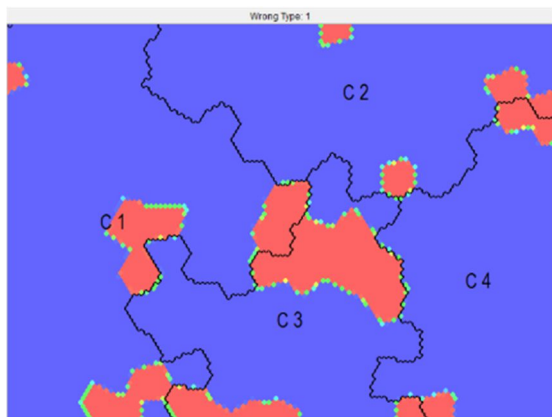


Figure 20. Wrong Type SOM

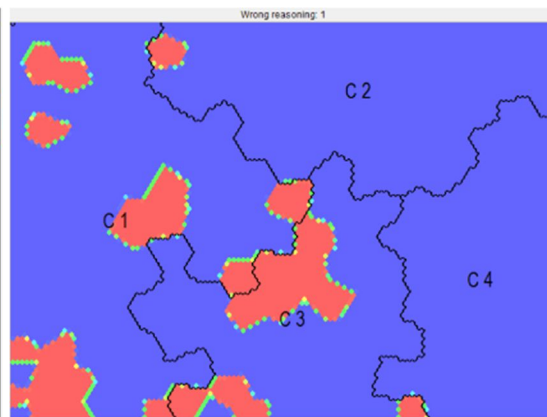


Figure 21. Wrong Reasoning SOM

4 Discussion

4.1 Clustering Interpretation

The application of the SOM algorithm resulted in the reorganisation of accidents originally arranged in a 238 x 53 matrix (Figure 4) into topographic maps (Figures 5 and 7-21). Having the accidents grouped by similarity (considering the contributing factors), it is now possible to identify common patterns and highlight key relationships within the dataset. Moreover, the results are presented through a graphical interface, allowing analysts to effectively see the most frequent features or have further insight into contributing factors which might be of interest (such as Design Failures, Human Erroneous Actions, Quality Control and Task Allocation).

The combination of three organisational factors (Inadequate Task Allocation, Inadequate Quality Control and Design Failure - Figures 8, 9 and 10) occupied most of the Cluster 1 area, meaning that these aspects are leading contributors to the grouping. Human erroneous actions contribute to 70% of the cluster, being the Wrong Place (Figure 12) the most relevant one, covering more than a half of the grouping. The analysis of the individual maps clearly shows that Inadequate Procedures (Figure 13) are highly associated with this specific type of human erroneous action, meaning that incorrect, incomplete, ambiguous or instructions open to interpretation provoked specific problems to implement a sequence of operational movements. A deep relationship between Inadequate Procedures (Figure 13) and Insufficient Knowledge (Figure 11) can be perceived in Cluster 1, denoting that written instructions presumed some level of specific knowledge to recognise the situation and complete the operation, which was not the case in many events.

An example of this type of accident was described in a US Chemical Safety and Hazard Investigation Board (2004) safety bulletin, when operators were assigned to a cleaning process in a petrochemicals plant. They executed a nitrogen gas purging exactly as required by the written procedures, in order to remove a hazardous mixture from the pipe. Afterwards, they started a steam purge to finish the service. However, the procedural steps were not sufficiently detailed to ensure the removal of the mixture from the pipe, especially in the low points, and failed to describe the consequences of having flushing liquid in the system. The operators were unaware of the possibility of having residues in the line, as well as the chemical reactions that could occur. The steam purge

heated the peroxide/alcohol mix above its thermal decomposition temperature, resulting in an explosion and fire. The SOM map exploration shows that this combination of contributing factors is not an isolated episode, but a recognisable pattern (or tendency) in Cluster 1, which should prompt the attention of risk analysts.

The correlation of further factors, such as the design failure which allowed an unnecessary low-point section in the pipe route and the failure of the quality control to identify the low-point trap as well as the deficient procedure are also persistent in this grouping.

As in Cluster 1, Cluster's 2 most common features were Design Failure accompanied by Inadequate Task Allocation and Inadequate Procedures. However, the results for these factors are close to the overall dataset figures (65.97%, 60.07% and 44.09%, respectively) and thus cannot be considered to be major influencing factors to generate this clustering. A noticeable feature in this cluster was Adverse Ambient Conditions, which attained 14% in this grouping. The exploration of these events demonstrates that not only major natural events such as hurricanes, typhoons or earthquakes should be considered from a risk and safety management perspective, but also more common adverse situations like torrential rain, electrical storms and even the presence of airborne particles. A straightforward example of the latter is the case where haze from forest fires carried atmospheric particles to the intake of an air separation unit of a gas processing facility, causing an explosion and a large fire.

Equipment Failure (as shown in Figure 14) dominates almost the whole area of Cluster 3, but, in sharp contrast with Cluster 1, the association with Maintenance Failure (Figure 15) is not relevant any longer. In fact, the analysis of the maps indicates that the equipment failure events tend to be associated with Design Failure and/or Insufficient Skills (Figure 16) for this grouping. Therefore, it can be learned that enhancing maintenance cannot be considered the only solution to minimise the possibility of equipment failures. The lack of skills (training / experience) to operate a system or equipment may well be combined with equipment faults, as well as with design shortcomings. 71.8% of the Cluster 3 area was covered by human erroneous actions, mostly Wrong Time and Wrong Type (Figures 18 and 20). A profounder analysis of specific cognitive functions influencing human actions can be also attained. Observation Missed, Wrong Reasoning and Inadequate Plan maps represented many cases where events or signals that were supposed to trigger an action were missed; the operator misinterpreted a given signal or cue – a deduction or induction error; or the mental plan/solution to solve an issue was incomplete or wrong. The airplane accident report mentioned in the introduction (*Bureau d'Enquêtes et d'Analyses pour la sécurité de l'aviation civile*, 2011) perfectly illustrates circumstances where an equipment failure due to a design shortcoming can trigger cognitive disturbances, leading to errors and ultimately major accidents. The lack of skills of the co-pilot to handle some problems (i.e. the approach to stall) at high altitudes made him build an erroneous mental plan to react to the undesirable situation. He adopted a tactic applicable to low altitudes, destabilising the flight path with inappropriate control inputs. Therefore, these tendencies reveal that specific training, aimed at dealing with critical conditions and major hazards, is very complex and must be carefully selected. Instead of reassuring written procedures or transmitting instructions, an effective training strategy must embrace the systemic development of a problem-solving mindset. Although most of the simulation and training strategies are focused on *conditioning* the human to expected or predictable scenarios, critical situations will demand advanced decision-

making skills and should focus on processes and techniques aimed at the identification and development of adequate operation alternatives.

Equipment Failure is also the main contributor to Cluster 4, but the accidents within this grouping presented different characteristics from the former cluster. The map's analysis for this cluster (i.e. Figures 9, 10 and 14) unmistakably shows that equipment problems were accompanied by quality control issues and design shortcomings. Although these relationships are quite clear, the low mode of 2 contributing factors prevents further tendencies from being accurately inferred in this grouping.

5 Conclusions

The successful conversion of multi-attribute, complex data from a major-accident dataset into a 2-D array revealed numerous possibilities of data clustering and interpretation, in order to disclose features, facilitate risk communication and enhance the learning process. The usage of graphical visualisation techniques such as topographic maps, which were generated by the SOM algorithm (Kohonen, 2001) in this research, provided additional means to help stakeholders absorb risk information and synchronise the textual explanation with meaningful visual representations.

The application of an artificial neural network approach permitted the identification of common patterns and comparable contributing factors within four major accidents groups, revealing interfaces and conveying information to operators, designers, risk managers and the general public.

Beyond the visual aid provided by the maps' construction, it was possible to directly correlate real accidents with images, creating and enhancing a full learning experience that can be further expanded, depending on the objectives and the targeted public. This is due to the data mining approach, which fully preserved the input data (the MATA-D 238 x 53 Matrix) in the output space (the 2-D Maps) and allowed the retrieval of the dataset records.

Also, the interpretation of the graphs can help to understand and communicate the relationship between contributing causes for major accidents. Figures 14 and 15, for instance, can be used to show to a sceptical operations manager that very advanced maintenance procedures do not guarantee that equipment will not fail, and further measures might be necessary.

The strategy of representing accident data in maps allows the fast transmission of relevant information and increases the possibility that stakeholders will fix and remember the lessons learned from accidents, minimising the dominance of biased concepts such as the oversimplification of addressing human errors as the main cause of major disasters.

6 Acknowledgements

The authors gratefully acknowledge the insights from Dr. Franz Knoll (NCK Inc.). This study was partially funded by CAPES [Grant nº 5959/13-6].

7 References

Andreev and Argyrou, 2011. Using Self-Organizing Map for Data Mining: A Synthesis with Accounting Applications. *Data Mining: Foundations and Intelligent Paradigms*, vol 3. Ed 1. Berlin: Springer.

Airbus, 2016. *Airbus Family Figures March 2016 Edition*. Blagnac Cedex: Airbus Print Centre. Available from: http://www.airbus.com/fileadmin/media_gallery/files/brochures_publications/aircraft_families/Airbus-Family-figures-booklet-March2016.pdf (Accessed: 04 May 2016).

BBC News, 2014. *South Korea ferry disaster: New arrests*. Available from: <http://www.bbc.co.uk/news/world-asia-27124528> (Accessed: 01 March 2016).

Bureau d'Enquêtes et d'Analyses pour la sécurité de l'aviation civile, 2011. *Final Report on the accident on 1st June 2009 to the Airbus A330-203* [Online]. Available from: <http://www.bea.aero/docspa/2009/f-cp090601.en/pdf/f-cp090601.en.pdf> (Accessed: 11 April 2014).

Cullen, W. (1990) *The public inquiry into the Piper Alpha disaster*. London: Her Majesty's Stationery Office (HMSO).

Dahle et al., 2012. Major accidents and their consequences for risk regulation, *In: Advances in Safety, Reliability and Risk Management*, Bérenguer, Grall & Guedes Soares (eds). London: Taylor & Francis Group.

Dekker, S., 2014. *Field Guide to Understanding 'Human Error'*. 3rd ed. Farnham : Ashgate Publishing Ltd, 2014.

Dhillon, B.S., 1986. *Human Reliability: With Human Factors*. New York: Pergamon Press Inc.

Dhillon, B.S., 2007. *Human reliability and error in transportation systems*. 1st ed. London: Springer-Verlag.

European Safety, Reliability and Data Association (ESReDA), 2015. Barriers to learning from incidents and accidents. Available from: <http://www.esreda.org/Portals/31/ESReDA-barriers-learning-accidents.pdf> (Accessed 08 April 2016).

Fukasawa, J., Okusaki, M., 2012. *Reform of the Nuclear Safety Regulatory Bodies In Japan*, International Nuclear Law Association Congress, 8–11 October, Manchester, UK.

Graeber, C., 1999. *The Role of Human Factors in Aviation Safety in Aero Magazine QTR_04 1999* (p. 23-31). The Seattle: Boeing Commercial Airplanes Group.

Hollnagel, E., 1993. The phenotype of erroneous actions. *International Journal of Man-Machine Studies*, 39, 1-32.

Hollnagel, E., 1998. *Cognitive Reliability and Error Analysis Method*. 1st ed. Oxford: Elsevier Science Ltd.

Hopkins, A., 2006. *A corporate dilemma: To be a learning organisation or to minimise liability. Technical report*, Australian National University, Canberra, Australia. National Center for OSH regulation Working Paper 43. Available from:

<https://digitalcollections.anu.edu.au/bitstream/1885/43147/2/wp43-corporatedilemma.pdf>.
(Accessed: 28 May 2016).

Ishikawa, M., 2015. *A Study of the Fukushima Daiichi Nuclear Accident Process - What caused the core melt and hydrogen explosion?* Tokyo: Springer.

Johnson, C., 2008. Ten contentions of corporate manslaughter legislation: Public policy and the legal response to workplace accidents. *Safety Science Journal* 46, 349-370.

Kohonen, T., 1996. Engineering Applications of the self-organizing map, *Proceedings of the IEEE* 10, Vol. 84, October 1996.

Kohonen, T., 1998. The *self-organizing map*. *Neurocomputing* 21, 1-6.

Kohonen, T., 2001. *Self-Organizing Maps*. 3rd ed. Berlin: Springer.

Kohonen, T., 2013. 'Essentials of the self-organizing map', *Neural Networks*, 37, pp. 52-65. Elsevier.

Kurokawa, K. et al., 2012. *The Official Report of The Fukushima Nuclear Accident Independent Investigation Commission - Executive Summary* [Online] Tokyo: The National Diet of Japan. Available from: https://www.nirs.org/fukushima/naiic_report.pdf (Accessed: 6 November 2014).

Leveson, N., 2004. 'A new accident model for engineering safer systems', *Safety Science Journal* 42, 237-270.

Leveson, N., 2011. 'Applying systems thinking to analyze and learn from events', *Safety Science Journal* 49, 55-64.

Moura, R. et al., 2016. 'Learning from major accidents to improve system design', *Safety Science Journal* 84, 37-45.

Paté-Cornel, M., 1993. 'Learning from the Piper Alpha Accident: A Postmortem Analysis of Technical and Organizational Factors', *Risk Analysis*, Vol. 13, No. 2, 1993 215-232.

Perrow, C., 1999. *Normal Accidents: Living With High-Risk Technologies*. New Jersey: Princeton University Press.

Reason, J., 1990. *Human Error*. Cambridge: Cambridge University Press.

Reason, J., 2000, 'The Freudian slip revisited', *Psychologist*, 13 (12), pp. 610-611.

Reason, J., 1997. *Managing the Risks of Organizational Accidents* 1st ed. Farnham: Ashgate Publishing Ltd.

Reason, J., 2013. *A Life in Error: From Little Slips to Big Disasters*. 1st ed. Farnham: Ashgate Publishing Ltd.

Schneider et al., 2012 *The World Nuclear Industry Status Report 2010–2011 - Nuclear Power in a Post-Fukushima World and 25 Years after the Chernobyl Accident*. Washington, D.C.: Worldwatch Institute. Available from: <http://www.worldnuclearreport.org/The-World-Nuclear-Industry-Status-51.html> (Accessed: 21 April 2016).

Swain, A. & Guttman, H., 1983. *NUREG/CR 1278 - Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications*. Albuquerque: Sandia National Laboratories.

Taleb, N., 2007. *The Black Swan: The Impact of the Highly Improbable*. 2nd Ed. York: Allen Lane.

Hollnagel, E., 1993. The phenotype of erroneous actions. *International Journal of Man-Machine Studies*, 39, 1-32.

U.S. Chemical Safety and Hazard Investigation Board., 2004. *Safety Bulletin No. 2004-03-B, July 2004*.

Woods, D. et al., 2010. *Behind Human Error*. 2nd ed. Farnham: Ashgate Publishing Ltd.